

Definition 1. Let G be a group. Then a *coset* is a subgroup H of G which is either a *left coset* of H , that is $xH = \{xh : h \in H\}$ for some x in G , or a *right coset* $Hx = \{hx : h \in H\}$ of the same.

Definition 2. Let polynomials $f_1(x), \dots, f_r$ in $\mathbf{F}_q[x]$ be non-zero. Then the *least common multiple* $\text{lcm}(f_1(x), \dots, f_r(x))$ of $f_1(x), \dots, f_r(x)$ is the monic polynomial of the lowest degree which is a multiple of all $f_i(x)$, $i = 1, \dots, r$.

Problem 1. Prove that for non-zero polynomials $f_1(x), \dots, f_r(x)$ in $\mathbf{F}_q[x]$,

$$\text{lcm}(f_1(x), \dots, f_r(x)) = \text{lcm}(\text{lcm}(f_1(x), \dots, f_{r-1}(x)), f_r(x))$$

Note 1. Let $f_1(x), \dots, f_r(x)$ in $\mathbf{F}_q[x]$ have the factorisations,

$$\begin{aligned} f_1(x) &= a_1 (p_1(x))^{e_{11}} \cdots (p_n(x))^{e_{1n}} \\ &\vdots \\ f_r(x) &= a_r (p_1(x))^{e_{r1}} \cdots (p_n(x))^{e_{rn}} \end{aligned}$$

where $a_1, \dots, a_r$ are in $\mathbf{F}_q^*$, $e_{ij} \geq 0$, and $p_i(x)$ are distinct monic irreducible polynomials over $\mathbf{F}_q$, then

$$\text{lcm}(f_1(x), \dots, f_r(x)) = (p_1(x))^{\max(e_{11}, \dots, e_{r1})} \cdots (p_n(x))^{\max(e_{1n}, \dots, e_{rn})}$$

Theorem 1. Let $f(x), f_1(x), \dots, f_r(x)$ be polynomials over $\mathbf{F}_q$. If $f(x)$ is divisible by every polynomial f_i , for $i = 1, \dots, r$, then $f(x)$ is also divisible by $\text{lcm}(f_1(x), \dots, f_r(x))$.

Proof. Consider first the case where there are only two different polynomials, $f_1(x)$ and $f_2(x)$. The prime components of $f_1(x)$ and $f_2(x)$ may be grouped into those which are unique among them and those which are shared. Since

$$f(x) = u_1(x)f_1(x) + r_1(x)$$

and

$$f(x) = u_2(x)f_2(x) + r_2(x)$$

it follows that $f(x)$ contains both of these two groups of primes. In other words,

$$f(x) = u(x) \text{lcm}(f_1(x)f_2(x)) + r(x)$$

Next, consider the case where there are more than two f_i 's. Suppose for $f(x)$, that

$$f(x) = u_r(x) \operatorname{lcm}(f_1(x), \dots, f_r(x))$$

Then if we let

$$f_c(x) = \operatorname{lcm}(f_1(x), \dots, f_r(x))$$

and if we introduce another polynomial $f_{r+1}(x)$ such that

$$f(x) = u_{r+1}f_{r+1} + r_{r+1}(x)$$

then following the same line of reasoning as the above we have,

$$\operatorname{lcm}(f_1(x), \dots, f_{r+1}(x)) | f(x)$$



Definition 3. A non-empty subset S of a ring R is called a *subring* of R if the elements of S form a ring with respect to the operations defined in R .

Theorem 2. Let R be a ring. Then a non-empty subset S of R is a subring if and only if S is closed under addition, multiplication, and the formation of additive inverse.

Proof. Since S is a subset of R , additive associativity, identity and commutativity are inherited to S from R . The existence of the inverse for each element s in S is certain provided that the formation of an additive inverse is guaranteed. And similarly in the case of multiplication, both associativeness and distributiveness hold once we know that S is closed under multiplication.

¶

Definition 4. Let R be a ring. We call an *ideal* in R a sub ring I having such property that for all i in I , then both xi and ix are also in I for every element x in R . Further, if I is a proper subset of R , then it is called a *proper ideal*. By *trivial ideal* one means either the *zero ideal* $\{0\}$ consisting of the zero element alone, or the ring R itself.

Note 2. The significance of the ideals in a ring is that they let us construct other rings from the first. The cosets of a ring R is a partition of R into equivalence sets, which are non-empty and disjoint, the union of which is the whole of the ring R .

Definition 5. Let R be a ring and I an ideal in it. Then two elements x and y in R are said to be *congruent modulo I* , denoted by

$$x \equiv y \pmod{I}$$

if $x - y$ is in I . Since there is only ideal, we may write this congruence as simply $x \equiv y$.

Note 3. Congruences can be added and multiplied as if they were ordinary equations. In other words, if $x_1 \equiv x_2$ and $y_1 \equiv y_2$, then

$$x_1 + y_1 \equiv x_2 + y_2$$

and

$$x_1 y_1 \equiv x_2 y_2$$

Definition 6. Let R be a ring and let x be an element of R . Then the *coset* $[x]$ containing x is the set of all elements y such that $y \equiv x$. Then,

$$\begin{aligned}[x] &= \{y : y \equiv x\} = \{y : y - x \in I\} \\ &= \{y : y - x \in i \text{ for some } i \in I\} \\ &= \{y : y = x + i \text{ for some } i \in I\} \\ &= \{x + i : i \in I\} = x + I\end{aligned}$$

Furthermore, $[x] = [x_1]$ means that $x \equiv x_1$, that is to say, $x - x_1$ is in I . Here x and x_1 are called *representatives* of the coset which contains them.

Definition 7. A *quotient ring*, aka *residue-class*-, *factor*-, or *difference ring*, is a ring having the form of a quotient A/i of a ring A and one of its ideal i . In other words, the quotient ring of R with respect to I the ring

$$R/I = \{x + I : x \in R\}$$

where

$$x + I = \{x + i : i \in I\}$$

is the coset of an element x in R , and where addition and multiplication are defined as,

$$[x] + [y] = [x + y]$$

and

$$[x] \cdot [y] = [xy]$$

Theorem 3. The zero element of R/I is $0 + I = I$, the negative of $x + I$ is $(-x) + I$. If R is commutative, then R/I is also commutative. If R has an identity 1 and a proper ideal I , then R/I has an identity $1 + I$.

Theorem 4. Let R be a ring and I an ideal of R . Then, for x and y in R ,

$$(x + I) + (y + I) = (x + y) + I$$

and

$$(x + I)(y + I) = xy + I$$

Proof. Let a and b be any two elements of the ideal I . Then,

$$(x + a) + (y + b) = x + a + y + b = (x + y) + (a + b) = (x + y) + p$$

where $p = a + b$ is in I . Further,

$$\begin{aligned} (x + a)(y + b) &= xy + bx + ay + ab \\ &= xy + c + d + e = xy + f \end{aligned}$$

where $c = bx$, $d = ay$, $e = ab$ and $f = c + d + e$ are all elements of I . ¶

Note 4. Theorem 4 and Note 3 show that the quotient ring R/I defined in Definitions 7 is independent of the choice of x and y in the cosets $x + I$ and $y + I$. In other words, the cosets $[x + y]$ and $[xy]$ resulted from addition and respectively multiplication in no ways depend on the particular representatives x and y chosen for the cosets $[x]$ and $[y]$ that go into them. This means that, if $x_1 \equiv x$ and $y_1 \equiv y$, then

$$[x_1 + y_1] = [x + y]$$

and

$$[x_1 y_1] = [xy]$$

or equivalently

$$x_1 + y_1 \equiv x + y$$

and

$$x_1 y_1 \equiv xy$$

Example 1. Some examples of quotient ring are $\mathbf{Z}_2 = \mathbf{Z}/2\mathbf{Z}$ and $\mathbf{Z}_6 = \mathbf{Z}/6\mathbf{Z}$.

Theorem 5. The polynomial ring $F[x]$ is a commutative ring with identity.

Proof. $F[x]$ is a ring over the field F since under addition it is closed, associative and commutative, and has 0 as the identity and the inverse $-f(x)$, where $f(x) \in F[x]$; and under multiplication it is associative, distributive and commutative, and has 1 as the identity. $\blacksquare$

Definition 8. Let R be a commutative ring with identity. Then for any a in R the *principal ideal* generated by a is

$$\langle a \rangle = aR = \{ar : r \in R\}$$

Further, R is called *principal ideal ring* if all its ideals are of this form.

Theorem 6. Let F be a field. Then the polynomial ring $F[x]$ is a principal ideal ring.

Proof. The polynomial ring $F[x]$ being a commutative ring with identity, it remains only to show that all its ideals are of the form

$$\langle a \rangle R = aR = \{ar : r \in R\}$$

where a is in R . Let I be an ideal of $F[x]$. If $I = 0$, then I is a principal ideal generated by 0. If $I \neq 0$, then choose $0 \neq f(x) \in I$ such that

$$\deg f \leq \deg g$$

for all non-zero $g(x)$ in I . Write

$$g(x) = q(x)f(x) + r(x)$$

If $\deg g < \deg f$, then $q = 0$ and $r = g$. On the other hand, if $n = \deg f \leq \deg g$, then either r is 0 or $\deg r < \deg f$.

Let

$$f(x) = a_0x^n + \cdots + a_n$$

and

$$g(x) = b_0x^m + \cdots + b_m$$

Then, with $a_0 \neq 0$,

$$g(x) = a_0^{-1}b_0x^{m-n}f(x) + g_1(x) \quad (1)$$

where $\deg g_1 \leq m - 1$. Then

$$g_1(x) = q_1(x)f(x) + r(x) \quad (2)$$

From this it follows that either $r = 0$ or $\deg r < \deg f$. From Equation's 1 and 2,

$$g(x) = q(x)f(x) + r(x)$$

where

$$q(x) = a_0^{-1}b_0x^{m-n} + q_1$$

is in $F[x]$. If $r \neq 0$, then $r(x)$ is in I and $\deg r < \deg f$, which contradicts our choice of $f(x)$. Therefore $g = qf$ and I is a principal ideal generated by $f(x)$.

¶

Definition 9. Let R be a commutative ring with identity. Then a non constant $f(x)$ in $R[x]$ is said to be *reducible* if, for some $g(x)$ and $h(x)$ in $R[x]$,

$$f(x) = g(x)h(x)$$

implies either $\deg g(x) = 0$ or $\deg h(x) = 0$. Otherwise $f(x)$ is said to be *reducible*.

Theorem 7. Let F be a field $f(x)$ in $F[x]$ an irreducible polynomial. Then $F[x]/\langle f(x) \rangle$ is a field.

Proof. Let I be the ideal $\langle f(x) \rangle$ of $F[x]$ generated by $f(x)$. If $I = F[x]$, then $f(x)$ has an inverse, that is $1 = f(x)g(x)$ for some $g(x)$ in $F[x]$. Then $f(x)$ is a constant polynomial, which contradicts our statement of the theorem. Therefore $F[x]/I$ has at least two elements, and $F[x]/I$ being a polynomial ring it is a commutative ring with identity.

Let $g \in F[x]$ and $g \notin I$. Then,

$$J = \{a(x)f(x) + b(x)g(x) : a(x), b(x) \in F[x]\}$$

is an ideal of $F[x]$ and there exists $h(x)$ in $F[x]$ such that $J = \langle h(x) \rangle$. But

$$f(x) = 1f(x) + 0g(x)$$

is in J , and thus $f(x) = a(x)h(x)$ for some $a(x)$ in $F[x]$. The polynomial $f(x)$ being irreducible, either $\deg h(x) = 0$ or $\deg a(x) = 0$. If the latter is the case, then $a(x)$ is a unit in $F[x]$, and then $h(x)$ is in I , hence $J = I$, and hence a contradiction since we began with g being in J but not in I . Therefore it must be the case that $h(x)$ is a unit in $F[x]$, hence J is a unit, and thus

$$1 = a(x)f(x) + b(x)g(x)$$

for some $a(x)$ and $b(x)$ in $F[x]$. And then

$$1 + I = I + b(x)g(x) = (b(x) + I)(g(x) + I)$$

Thus $g(x) + I$ has an inverse and $F[x]/I$ is a field. ¶

Definition 10. Let K be a field and F a subfield of K . Then K is called an *extension* of the field F , denoted by

$$K|_F$$

Since K has multiplication, it is a vector space over F . The dimension of the vector space K over F is called the *degree*

$$[K : F]$$

of the extension K of F . The extension $K|_F$ is said to be *finite* if the degree $[K : F]$ is finite.

Definition 11. A *prime subfield* of a field F is the intersection of all subfields of F . It is the smallest of all subfields of F , and is unique. A *prime field* is a field which has no proper subfields.

Definition 12. Let $K|_F$ be an extension of a field F . Then $\alpha \in K$ is said to be *algebraic* over F if there exists $f(x)$ in $F[x]$ which has α as a root. Let α in K be algebraic over F and consider

$$A = \{f(x) \in F[x] : f(\alpha) = 0\}$$

Here A is an ideal of the principal ideal domain $F[x]$. Let $m_1(x)$ in $F[x]$ be a generator of A . If a is the coefficient of the highest power of x in $m_1(x)$, then

$$m(x) = a^{-1}m_1(x)$$

is a monic polynomial with $\deg m(x) = \deg m_1(x)$, and $m(x)$ is also a generator of A . Let

$$m(x) = r(x)s(x)$$

for some $r(x)$ and $s(x)$ in $F[x]$. Then either $r(\alpha) = 0$ or $s(\alpha) = 0$, that is either $m(x)|r(x)$ or $m(x)|s(x)$. But $\deg m = \deg r + \deg s$, therefore either $\deg r(x) = 0$ or $\deg s(x) = 0$. Hence $m(x)$ is irreducible. Since $m(x)$ is monic, irreducible and is of the least degree possible while admitting α as a root, therefore $m(x)$ is called the *minimal polynomial* of α over $F[x]$.

Theorem 8. Let C be an (n, k) linear code over F_q with prith-check matrix H , and $d(C)$ the smallest number of column of H that are linearly dependent. Then if every subset of $2t$ or fewer columns of H is linearly independent, the code is capable of correcting all error patterns of weight $w \leq t$.

Proof. When $q = 2$, linear independence amounts to summing to $\mathbf{0}$. The code words of C are those vectors $\mathbf{x}$ in $V_n(F_q)$ for which

$$H\mathbf{x}^T = \mathbf{0}$$

But $H\mathbf{x}^T$ is a linear combination of the columns of H , that is to say, if

$$H = [\mathbf{c}_1 \quad \cdots \quad \mathbf{c}_n]$$

then

$$H\mathbf{x}^T = x_1\mathbf{c}_1 + \cdots + x_n\mathbf{c}_n$$

Hence a non-zero code word of weight w gives a nontrivial linear dependence among w columns of H , and vice versa. $\blacksquare$

Corollary 8[1]. If $q = 2$ and all possible linear combinations of up to e columns are distinct, then

$$d(C) \geq 2e + 1$$

and C can then correct all patterns of weight e or less.

Note 5. Hamming codes correct single errors. An extension of this is to the Bose-Chaudhuri-Hocquenghem codes which could correct multiple errors. In the case of Hamming code of length $n = 2^m - 1$, the parity-check matrix is given by

$$H = [\mathbf{v}_0 \quad \cdots \quad \mathbf{v}_{n-1}]$$

where $(\mathbf{v}_0 \quad \cdots \quad \mathbf{v}_{n-1})$ is some ordering of the $2^m - 1$ non-zero column vectors in $V_m = V_m(F_2)$. The $m \times n$ matrix H takes m parity-check bits for the code to be able to correct one error. We may extend H such that it has m more rows and could correct two errors. Then,

$$H_2 = \begin{bmatrix} \mathbf{v}_0 & \cdots & \mathbf{v}_{n-1} \\ \mathbf{w}_0 & \cdots & \mathbf{w}_{n-1} \end{bmatrix}$$

where $\mathbf{w}_0, \dots, \mathbf{w}_{n-1}$ are in V_m .

Since $\mathbf{v}_i$'s are distinct, we may look at the mapping from $\mathbf{v}_i$ to $\mathbf{w}_i$ as a function from V_m into itself, then

$$H_2 = \begin{bmatrix} \mathbf{v}_0 & \cdots & \mathbf{v}_{n-1} \\ \mathbf{f}(\mathbf{v}_0) & \cdots & \mathbf{f}(\mathbf{v}_{n-1}) \end{bmatrix}$$

Then H_2 will define a code which corrects two errors if and only if the syndromes of the $1 + n + \binom{n}{2}$ error patterns of weights 0, 1 and 2 are all distinct. Any such syndrome is a sum of a subset of columns of H_2 , and therefore a vector in V_{2m} . Let the syndrome be $\mathbf{s} = (s_1 \ \dots \ s_{2m}) = (\mathbf{s}_1 \ \mathbf{s}_2)$, where $\mathbf{s}_1 = (s_1, \dots, s_m)$ and $\mathbf{s}_2 = (s_{m+1}, \dots, s_{2m})$ are both in V_m . Defining $\mathbf{f}(\mathbf{0}, \mathbf{0}) = \mathbf{0}$ we consider a pair of errors occurring at i^{th} - and j^{th} position's, $\mathbf{s} = (\mathbf{v}_i + \mathbf{v}_j, \mathbf{f}(\mathbf{v}_i) + \mathbf{f}(\mathbf{v}_j))$. Then the system of equations,

$$\begin{aligned} \mathbf{u} + \mathbf{v} &= \mathbf{s}_1 \\ \mathbf{f}(\mathbf{u}) + \mathbf{f}(\mathbf{v}) &= \mathbf{s}_2 \end{aligned}$$

has at most one solution $(\mathbf{u}, \mathbf{v})$ for each pair of vectors from V_m .

By trial and error we may find neither the linear mapping $\mathbf{f}(\mathbf{v}) = T\mathbf{v}$ nor the nonlinear polynomial of degree 2 works, but $\mathbf{f}(\mathbf{v}) = \mathbf{v}^3$ does. The matrix

$$H_2 = \begin{bmatrix} \alpha_0 & \cdots & \alpha_{n-1} \\ \alpha_0^3 & \cdots & \alpha_{n-1}^3 \end{bmatrix}$$

is the parity-check matrix of a binary code of length $n = 2^m - 1$ which corrects up to two errors. A vector $\mathbf{c} = (c_0 \cdots c_{n-1})$ in $V_n(F_2)$ is a code word in the code defined by H_2 if and only if

$$\sum_{i=0}^n c_i \alpha_i = \sum_{i=0}^n c_i \alpha_i^3 = 0$$

Since the $2m$ rows of the matrix H_2 over F_2 may not be all linearly independent, the dimension of the code is

$$d(C) \geq n - 2m = 2^m - 1 - 2m$$

Definition 13. The *Vandermonde matrix* is defined as

$$A = \begin{bmatrix} 1 & \cdots & 1 \\ a_1 & \cdots & a_r \\ \vdots & \ddots & \vdots \\ a_1^{r-1} & \cdots & a_r^{r-1} \end{bmatrix}$$

Theorem 9. Let $a_1, \dots, a_r$ be distinct non-zero elements of a field. Then the Vandermonde matrix is such that

$$\begin{vmatrix} 1 & \cdots & 1 \\ a_1 & \cdots & a_r \\ \vdots & \ddots & \vdots \\ a_1^{r-1} & \cdots & a_r^{r-1} \end{vmatrix} \neq 0$$

Proof. Subtracting $\text{row}(i + 1) - a_1 \text{row } i$, $i = 1, \dots, r - 1$, yields,

$$\begin{aligned}
 \det A &= \begin{vmatrix} 1 & 1 & \cdots & 1 \\ 0 & a_2 - a_1 & \cdots & a_r - a_1 \\ 0 & a_2(a_2 - a_1) & \cdots & a_r(a_r - a_1) \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_2^{r-2}(a_2 - a_1) & \cdots & a_r^{r-2}(a_r - a_1) \end{vmatrix} \\
 &= (a_2 - a_1) \cdots (a_r - a_1) \begin{vmatrix} 1 & \cdots & 1 \\ a_2 & \cdots & a_r \\ \vdots & \ddots & \vdots \\ a_2^{r-2} & \cdots & a_r^{r-2} \end{vmatrix} \\
 &= (a_2 - a_1) \cdots (a_r - a_1) \cdot (a_3 - a_2) \cdots (a_r - a_2) \begin{vmatrix} 1 & \cdots & 1 \\ a_3 & \cdots & a_r \\ \vdots & \ddots & \vdots \\ a_3^{r-3} & \cdots & a_r^{r-3} \end{vmatrix} \\
 &\quad \vdots \\
 &= \prod_{i>j} (a_i - a_j)
 \end{aligned}$$

Then, since a_i are distinct and non-zero, therefore $\det A$ is non-zero. ¶

Theorem 10. Any square matrix having a non-zero determinant has all its columns linearly independent.

Proof. Let A be an $r \times r$ matrix, and that $|A| \neq 0$. Then suppose the columns of A are linearly dependent. Then one may write some column of A as a linear combination of the others, for example

$$\mathbf{c}_j = \sum_{\substack{i=1 \\ i \neq j}}^r a_i \mathbf{c}_i$$

Then if column $\mathbf{c}_j$ is replaced by $\mathbf{c}_j - \sum_{\substack{i=1 \\ i \neq j}}^r a_i \mathbf{c}_i$ gives a matrix B with $|B| = |A|$. But B also has a column whose all elements are zeros, which means that

$$|A| = |B| = 0$$

a contradiction and thus the proof. ¶

Theorem 11. Let $(\alpha_0, \dots, \alpha_{n-1})$ be an ordering of non-zero elements of $\mathbf{F}_{2^m}$, and let t be a positive integer such that

$$t \leq 2^{m-1} - 1$$

Then the matrix

$$H = \begin{bmatrix} \alpha_0 & \cdots & \alpha_{n-1} \\ \alpha_0^3 & \cdots & \alpha_{n-1}^3 \\ \vdots & \ddots & \vdots \\ \alpha_0^{2t-1} & \cdots & \alpha_{n-1}^{2t-1} \end{bmatrix}$$

is the parity-check matrix of a binary (n, k) -code capable of correcting all error patterns of weight $w \leq t$, with dimension

$$k \geq n - mt$$

Proof. A vector $\mathbf{c} = (c_0, \dots, c_{n-1})$ in $V_n(F_2)$ is a code word if and only if $H\mathbf{c}^T = \mathbf{0}$. Thus,

$$\sum_{i=0}^{n-1} c_i \alpha_i^j = 0$$

for $j = 1, 3, \dots, 2t - 1$. We simplify this by using the fact that $(x + y)^2 = x^2 + y^2$ in characteristic 2, and $x^2 = x$ in F_2 . Hence,

$$\left(\sum_{i=0}^{n-1} c_i \alpha_i^j \right)^2 = \sum_{i=0}^{n-1} c_i^2 \alpha_i^{2j} = \sum_{i=0}^{n-1} c_i \alpha_i^{2j}$$

for $j = 1, 3, \dots, 2t - 1$, which gives us

$$\sum_{i=0}^{n-1} c_i \alpha_i^j$$

for $j = 1, 2, \dots, 2t$.

Therefore we could also use the parity-check matrix

$$H' = \begin{bmatrix} \alpha_0 & \cdots & \alpha_{n-1} \\ \alpha_0^2 & \cdots & \alpha_{n-1}^2 \\ \vdots & \ddots & \vdots \\ \alpha_0^{2t} & \cdots & \alpha_{n-1}^{2t} \end{bmatrix}$$

According to Theorem 8 H' is a parity-check matrix which corrects t errors if and only if every subset of $2t$ or fewer columns of H' is linearly independent. Next, since a subset of $r \leq 2t$ columns of H' has the form

$$A = \begin{bmatrix} a_1 & \cdots & a_r \\ a_1^2 & \cdots & a_r^2 \\ \vdots & \ddots & \vdots \\ a_1^{2t} & \cdots & a_r^{2t} \end{bmatrix}$$

where $a_1, \dots, a_r$ are distinct non-zero elements of F_{2m} , we may consider the matrix

$$A' = \begin{bmatrix} a_1 & \cdots & a_r \\ \vdots & \ddots & \vdots \\ a_1^r & \cdots & a_r^r \end{bmatrix}$$

which is nonsingular since its determinant by the Vandermonde determinant theorem, Theorem 9, is

$$\det A' = a_1 \cdots a_r \begin{vmatrix} 1 & \cdots & 1 \\ a_1 & \cdots & a_r \\ \vdots & \ddots & \vdots \\ a_1^{r-1} & \cdots & a_r^{r-1} \end{vmatrix} = a_1 \cdots a_r \prod_{i < j} (a_j - a_i) \neq 0$$

Then the columns of A' , and hence those of A , cannot be linearly dependent, and therefore the code corrects all error patterns of weight up to t . Now H , as a matrix with entries from F_2 rather than F_{2^m} , has dimensions $mt \times n$, hence the dual code has dimension $k \leq mt$, and the code has dimension $k \geq n - mt$. $\blacksquare$

Theorem 12. Let C be a linear (n, k) -code over $GF(q)$ with parity-check matrix H . Then the minimum distance of C is d if and only if any $d-1$ columns of H are linearly independent but some d columns are linearly dependent.

Proof. The minimum distance of a code $d(C)$ is equal to the minimum of the weights of the non-zero code words. Let $\mathbf{x} = x_1 \cdots x_n$ be a vector in $V(n, q)$. Then $\mathbf{x}$ is in C if and only if

$$\mathbf{x}H^T = \mathbf{0}$$

if and only if

$$x_1 \mathbf{h}_1 + \cdots + x_n \mathbf{h}_n = \mathbf{0}$$

where $\mathbf{h}_1, \dots, \mathbf{h}_n$ are the columns of H . Therefore there is a set of d linearly dependent columns of H corresponding to each code word $\mathbf{x}$ of weight d . On the other hand, if there existed a set of $d-1$ linearly dependent columns of H , then there would exist some scalars $x_{i_1}, \dots, x_{i_{d-1}}$, not all zero, such that $\sum_{j=1}^{d-1} x_{i_j} \mathbf{h}_{i_j} = \mathbf{0}$. But if this were the case, then $\mathbf{x}H^T = \mathbf{0}$ and so would be a code word of weight $0 < d < d(C)$. $\blacksquare$

Theorem 13. The maximum dictionary size m such that there exists a q -ary (n, m, d) -code is

$$A_q(n, d) \leq q^{n-d+1}$$

Proof. Let C be a q -ary (n, m, d) -code. If we remove the last $d - 1$ coordinates from each code word, then the m vectors of length $n - d + 1$ so obtained must be distinct, otherwise $d(C)$ must be less than d , which would contradict the statement above. Therefore $m \leq q^{n-d+1}$.

Theorem 14. Let C be the code over $GF(q)$, where q is a prime number, and C is defined to have the parity-check matrix

$$H = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 3 & \cdots & n \\ 1 & 2^2 & 3^2 & \cdots & n^2 \\ \vdots & & & \ddots & \vdots \\ 1 & 2^{d-2} & 3^{d-2} & \cdots & n^{d-2} \end{bmatrix}$$

where $d \leq n \leq q - 1$. If q is a prime-power, then

$$A_q(n, d) = q^{n-d+1}$$

Proof. We have,

$$C = \left\{ x_1 \cdots x_n \in V(n, q) \text{ s.t. } \sum_{i=1}^n i^j x_i = 0 \text{ for } j = 0, 1, \dots, d-2 \right\}$$

Any $d-1$ columns form a Vandermonde matrix, and therefore by Theorem's 9 and 10 are linearly independent. By Theorem 12 C has a minimum distance d and therefore is a q -ary (n, q^{n-d+1}, d) -code. The proof follows since C meets the Singleton bound of Theorem 13. $\blacksquare$

Problem 2. Find the decoding procedure for the BCH codes.

Solution. Assume that $d = 2t + 1$ and H has $2t$ rows. Suppose the code word $\mathbf{c} = c_1 \cdots c_n$ is transmitted and the vector $\mathbf{r} = r_1 \cdots r_n$ is received. Assuming that at most t errors have occurred, let $x_1, \dots, x_t$ be their positions and $m_1, \dots, m_t$ their respective magnitudes. Then the syndrome is

$$(s_1, \dots, s_{2t}) = \mathbf{r}H^T$$

and we have

$$s_j = \sum_{i=1}^n r_i i^{j-1} = \sum_{i=1}^t m_i x_i^{j-1} \quad (3)$$

for $j = 1, \dots, 2t$.

Solution. Then from

$$\phi(\theta) = \frac{m_1}{1 - x_1\theta} + \frac{m_2}{1 - x_2\theta} + \cdots + \frac{m_t}{1 - x_t\theta} \quad (4)$$

and

$$\frac{m_i}{1 - x_i\theta} = m_i (1 + x_i\theta + x_i^2\theta^2 \cdots)$$

together with Equation 3, we have

$$\phi(\theta) = s_1 + s_2\theta + \cdots + s_{2t}\theta^{2t-1} + \cdots$$

Also, from Equation 4 we have

$$\phi(\theta) = \frac{a_1 + a_2\theta + a_3\theta^2 + \cdots + a_t\theta^{t-1}}{1 + b_1\theta + b_2\theta^2 + \cdots + b^t\theta^t} \quad (5)$$

Hence,

$$(s_1 + s_2\theta + s_3\theta^2 + \cdots) (1 + b_1\theta + b_2\theta^2 + \cdots + b_t\theta^t) = a_1 + a_2\theta + \cdots + a_t\theta^{t-1}$$

Solution. Which gives us

$$a_1 = s_1 \quad \text{and} \quad a_i = \sum_{j=0}^{i-1} s_{i-j} b_j, \quad i = 2, \dots, t \quad (6)$$

and

$$0 = \sum_{j=0}^t s_{i-j} b_j, \quad i = t+1, \dots, 2t \quad (7)$$

With a_i and b_i known we may turn Equation 5 into partial fractions

$$\phi(\theta) = \frac{p_1}{1 - q_1\theta} + \dots + \frac{p_t}{1 - q_t\theta}$$

and therefore $m_i = p_i$ and $x_i = q_i$, for $i = 1, \dots, t$, and the system in Equation 3 is solved. Algorithm 1 then gives the procedure for error correction.

#

Note 6. The polynomial

$$\sigma(\theta) = 1 + b_1\theta + b_2\theta^2 + \cdots + b_t\theta^t = (1 - x_1\theta) \cdots (1 - x_t\theta) \quad (8)$$

can be used to locate the location of the errors. The polynomial

$$\omega(\theta) = a_1 + a_2\theta + \cdots + a_t\theta^{t-1}$$

can be used to find the magnitude of the errors.

Algorithm 1 *Procedure for correcting up to t errors in BCH codes.*

input: $\mathbf{r}$

find $s_1, \dots, s_{2t}$

$e \leftarrow$ maximum number of equations in Equation 7

for $i = e + 1$ to t **do**

$b_i \leftarrow 0$

endfor

$(b_1, \dots, b_e) \leftarrow$ **solve** the first e equations of Equation 7

$(z_1, \dots, z_e) \leftarrow$ **find** the e zeros of Equation 8

$(a_1, \dots, a_e) \leftarrow$ **solve** Equation 6

for $i = 1$ to e **do**

$$m_i \leftarrow \frac{a_1 + a_2 x_i + \dots + a_e x_i^{e-1}}{\prod_{\substack{j=1 \\ j \neq i}}^e (1 + x_j x_i)}$$

endfor